# Bayesian model averaging

Μ. Παρζακώνης

ΠΜΣ "Εφαρμοσμένη Στατιστική"

2009

# Σύνοψη

1. Στατιστική πρακτική
   - Επιλογή βάση Κριτηρίων Πληροφορίας
   - Προβλήματα διαδικασιών επιλογής
2. BMA...
   - Η βασική ιδέα
   - Βασικές Υλοποιήσεις
   - Αλγόριθμοι υλοποίησης
3. Εφαρμογή
4. Περαιτέρω θέματα στη BMA
   - Εναλλακτικός προσδιορισμός των $w_i$
   - Το $w$ ως τυχαία μεταβλητή
   - Εμπειρική εκτίμηση του $\phi$
   - Η κατανομή των Zellner-Siow

Δοθέντων δεδομένων $X = x$ θέλουμε να προβλέψουμε/εκτιμήσουμε κ.α μια ποσότητα, έστω $\omega$ .

Η συνήθης διαδικασία βασιζεται στην καταγραφή των υποψήφιων μοντέλων και την επιλογή του επικρατέστερου με βαση (κάποιο) κρίτηριο.

# Επιλογή βάση Κριτηρίων Πληροφορίας

## Ορισμός

Η ποσότητα $IC_n(q_\gamma) = -2logL + \alpha n$, $\alpha > 0$ είναι ένα Γενικευμένο Κριτήριο Πληροφορίας (GIC) αν και μονο αν $\lim_{n\to\infty} \frac{\alpha}{n} = 0$.

## Γνωστά Κριτήρια Πληροφορίας

| $IC$ | $\alpha$ |
|---|---|
| $AIC$ | 2 |
| $BIC$ | $\log n$ |
| $HC(c)$ | $c \log\log n$ |

### Προβλήματα επιλογής μοντέλου

Αυξημένο σφάλμα Τύπου I/Εσφαλμένα "στενά" δ.ε.

### Μερικά αποτελέσματα σχετικά με τα Κριτήρια Πληροφορίας...

- (AIC) Είναι αποδοτικό ⇒ Αριθμός αναμενόμενων μεταβλητών «λανθασμένης» επιλογής $> 0$.
- (BIC) Είναι συνεπές ⇒ Αριθμός αναμενόμενων μεταβλητών «σωστής» επιλογής $< 0$. (Claeskens & Hjort, 2008)

Το $HC(c)$, $c > 1$ ελαχιστοποιεί τον αριθμό αναμενόμενων μεταβλητών «σωστής» επιλογής τις οποιες παραλείπει να επιλεξει το BIC.

# Προβλήματα διαδικασιών επιλογής

**Προβλήματα επιλογής μοντέλου**

Αυξημένο σφάλμα Τύπου I/Εσφαλμένα "στενά" δ.ε.

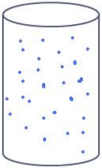
**Μερικά αποτελέσματα σχετικά με τα Κριτήρια Πληροφορίας...**

- (AIC) Είναι αποδοτικό $\Rightarrow$ Αριθμός αναμενόμενων μεταβλητών «λανθασμένης» επιλογής $> 0$.
- (BIC) Είναι συνεπές $\Rightarrow$ Αριθμός αναμενόμενων μεταβλητών «σωστής» επιλογής $< 0$. (Claeskens & Hjort, 2008)

Το $HC(c)$, $c > 1$ ελαχιστοποιεί τον αριθμό αναμενόμενων μεταβλητών «σωστής» επιλογής τις οποιες παραλείπει να επιλεξει το BIC.

## ...

Χρήση περισσότερων μοντέλων για συμπερασματολογία.



## Η πρωταρχική σχέση...

$$E(\omega|y) = \sum_{\gamma} E(\omega|y,\gamma)\pi(\gamma|y).$$

- *Βέλτιστη εκτίμηση (ως προς το ΜΤΣ).*
- *Βέλτιστη πρόβλεψη ως προς το λογαριθμικό κανόνα του Good.*

## Κλασσικό γραμμικό μοντέλο

$$Y = \alpha + X\beta + \varepsilon, \quad \varepsilon \sim N_n(0, \sigma^2 I_n).$$

Μπεϋζιανή προσθήκη...

$$\begin{aligned}(\alpha, \beta, \sigma^2) &\sim p(\alpha, \beta, \sigma^2),\\ \gamma &\sim \pi(\gamma).\end{aligned}$$

## Συνήθης μοντελοποίηση

- $\pi(\gamma)$, π.χ Διωνυμική,Poisson
- $p(\alpha, \beta, |\sigma^2, \gamma)$, κανονική
- $p(\sigma^2, |\gamma)$, αντίστροφη γάμμα

# Το (Μπεϋζιανό) Γραμμικό Μοντέλο

## Κλασσικό γραμμικό μοντέλο

$$Y = \alpha + X\beta + \varepsilon, \quad \varepsilon \sim N_n(0, \sigma^2 I_n).$$

Μπεϋζιανή προσθήκη...

$$\begin{aligned} (\alpha, \beta, \sigma^2) &\sim p(\alpha, \beta, \sigma^2), \\ \gamma &\sim \pi(\gamma). \end{aligned}$$

## Συνήθης μοντελοποίηση

- $\pi(\gamma)$, π.χ Διωνυμική,Poisson
- $p(\alpha, \beta, |\sigma^2, \gamma)$, κανονική
- $p(\sigma^2, |\gamma)$, αντίστροφη γάμμα

## Hoeting, Raftery & Madigan (1997)

- $\pi(\gamma) = 2^{-p}$,
- $p(\theta|\sigma^2, \gamma) \sim N_{q_\gamma+1}(\tilde{\theta}_\gamma, \sigma^2\Sigma^{-1}), \theta = (\alpha, \beta)$
- $p(\sigma^2|\gamma) \sim IG(\nu/2, \nu\lambda/2)$

*Εκ των προτέρων επιλογές*

- $\tilde{\theta}_\gamma = (\bar{y}, 0, 0, \ldots, 0)$,
- $\Sigma = \left(diag\{s_Y^{-2}, \frac{s_{X_1}^2}{\phi}, \frac{s_{X_2}^2}{\phi}, \ldots, \frac{s_{X_p}^2}{\phi}\}\right)^{-1}$,
- $(\nu, \lambda, \phi) = (2.58, 0.28, 2.85)$.

## Zellner

- $\pi(\gamma) = 2^{-p}$,
- $p(\beta|\sigma^2,\gamma) \sim N_{q_\gamma+1}(\tilde{\beta}_\gamma, \sigma^2\Sigma^{-1})$,
- $p(\alpha,\sigma^2|\gamma) \propto \sigma^{-2}$.

*Εκ των προτέρων επιλογές*

- $\tilde{\theta}_\gamma = 0$,
- $\Sigma = \phi\left(X^T X\right)^{-1}$,
- $\phi = n,\, p^2,\, max(n, p^2)$.

**Εκ των υστέρων κατανομές**

*Συζυγία...*

## Zellner

- $\pi(\gamma) = 2^{-p}$,
- $p(\beta|\sigma^2, \gamma) \sim N_{q_\gamma+1}(\tilde{\beta}_\gamma, \sigma^2\Sigma^{-1})$,
- $p(\alpha, \sigma^2|\gamma) \propto \sigma^{-2}$.

*Εκ των προτέρων επιλογές*

- $\tilde{\theta}_\gamma = 0$,
- $\Sigma = \phi \left(X^T X\right)^{-1}$,
- $\phi = n,\ p^2,\ max(n, p^2)$.

## Εκ των υστέρων κατανομές

*Συζυγία...*

1. Στατιστική πρακτική
   - Επιλογή βάση Κριτηρίων Πληροφορίας
   - Προβλήματα διαδικασιών επιλογής
2. BMA...
   - Η βασική ιδέα
   - Βασικές Υλοποιήσεις
   - Αλγόριθμοι υλοποίησης
3. Εφαρμογή
4. Περαιτέρω θέματα στη BMA
   - Εναλλακτικός προσδιορισμός των $w_i$
   - Το $w$ ως τυχαία μεταβλητή
   - Εμπειρική εκτίμηση του $\phi$
   - Η κατανομή των Zellner-Siow

# Αλγόριθμοι υλοποίησης

**Γεγονός (ανασταλτικό)...**

Ο αριθμός των υποψήφιων μοντέλων (συνήθως) είναι σημαντικός (εάν οχι τεράστιος), $2^p$ .

**Συνέπειες...**

- Δεν είναι δυνατή η εξαντλητική ανάλυση του συνόλου τους.
- Δεν είναι "εύκολος" ο υπολογισμός των εκ των υστέρων πιθανοτήτων.

# Αντιμετώπιση του προβλήματος

- Προσέγιση της εκ των υστέρων πιθανοφάνειας.
- Ελάττωση του συνόλου των υποψηφιων μοντέλων.

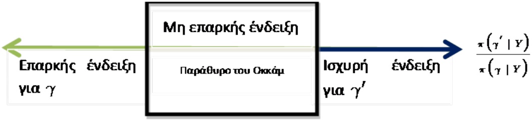
*Βήμα 1.* Βρες το μοντέλο με τη μέγιστη εκ των υστέρων πιθανότητα.

*Βήμα 2.* Σχημάτισε το σύνολο $\mathscr{A} = \{\gamma : \frac{\max_{\gamma} \pi(\gamma|y)}{\pi(\gamma|y)} \leq u\}$ για ένα $u$ τέτοιο ώστε το σύνολο να μην είναι μικρό ή υπερβολικά μεγάλο.

*Βήμα 3.* Σχημάτισε το σύνολο

$$\mathscr{B}=\left\{\gamma' : \exists \gamma \epsilon \Gamma, \gamma \Subset \gamma', \frac{\pi(\gamma|y)}{\pi(\gamma'|y)} > 1\right\} .$$

*Βήμα 4.* Το ζητούμενο σύνολο είναι το $\Gamma' = \mathscr{B} - \mathscr{A}$.




## Εργαλεία...

- leaps & bounds
- BIC

## Βήματα BIC...

Βήμα 2

$$BIC_{\gamma} \leq BIC_{\gamma *} + 2 \log u.$$

Βήμα 3

$$BIC_{\gamma'} \leq BIC_{\gamma}.$$

Έστω ότι βρισκόμαστε στο μοντέλο $\gamma$ , με παραμετρικό διάνυσμα $\theta_\gamma$.

*Βήμα 1.* Πρότεινε τη μετακίνηση στο μοντέλο $\gamma'$ με πιθανότητα $j(\gamma'|\gamma)$.

*Βήμα 2.* Πάρε ένα διάνυσμα $u$ από την $\varphi(u|\theta_\gamma,\gamma',\gamma)$ διάστασης $|\dim(\gamma') - \dim(\gamma)|$.

*Βήμα 3.* Θέσε $(\theta_{\gamma'}, u') = h_{\gamma,\gamma'}(\theta_\gamma, u)$, για $h_{\gamma,\gamma'}(\bullet,\bullet)$ μια αντιστρεψιμη συναρτηση. Τότε $\dim(\gamma') + \dim(u') = \dim(\gamma) + \dim(u)$.

*Βήμα 4.* Αποδέξου την μετακίνηση με πιθανότητα

$$\alpha = 1 \wedge \frac{p(y|\theta_{\gamma'},\gamma')\pi(\theta_{\gamma'}|\gamma')\pi(\gamma')j(\gamma|\gamma')}{p(y|\theta_\gamma,\gamma)\pi(\theta_\gamma|\gamma)\pi(\gamma)j(\gamma'|\gamma)} \frac{\varphi(u'|\theta_{\gamma'},\gamma',\gamma)}{\varphi(u|\theta_\gamma,\gamma',\gamma')} \frac{\partial h_{\gamma,\gamma'}(\theta_\gamma,u)}{d(\theta_\gamma,u)}$$

## Για τα μοντέλα HRM & Zellner

- Χρήση της $p(\theta_\gamma|y,\gamma)$ ως κατανομής πρότασης, $\phi(u|\theta_\gamma,\gamma',\gamma)$.
- Υιοθέτηση συμμετρικής κατανομής μετάβασης ($j(\gamma'|\gamma) = j(\gamma|\gamma')$).
- Ομοιόμορφη κατανομή στο Γ.

$$\alpha = 1 \wedge \frac{\pi(\gamma'|y)}{\pi(\gamma|y)}$$

## Ειδικοί αλγόριθμοι

- In-Out : $MCMCMC$ ή $MC^3$
- In-Out & Swap : $rJ - MCMC$

# Απλοποιήσεις του αλγόριθμου

## Για τα μοντέλα HRM & Zellner

- Χρήση της $p(\theta_\gamma|y,\gamma)$ ως κατανομής πρότασης, $\phi(u|\theta_\gamma,\gamma',\gamma)$.
- Υιοθέτηση συμμετρικής κατανομής μετάβασης ($j(\gamma'|\gamma) = j(\gamma|\gamma')$).
- Ομοιόμορφη κατανομή στο Γ.

$$\alpha = 1 \wedge \frac{\pi(\gamma'|y)}{\pi(\gamma|y)}$$

## Ειδικοί αλγόριθμοι

- In-Out : $MCMCMC$ ή $MC^3$
- In-Out & Swap : $rJ - MCMC$

## Το πρόβλημα...

Ο Ehrlich (1973) θεώρησε ότι η παραβατική συμπεριφορά είναι αποτέλεσμα λογικής οικονομικής απόφασης καθώς και ότι η πιθανότητα τιμωρίας αυτής της απόφασης δρα ανασταλτικά ως προς τη λήψη μιας τέτοιας απόφασης. Το σύνολο δεδομένων που μελετησε αποτελείται από τα δεδομένα 47 πολιτειών των Η.Π.Α το έτος 1960. Η μεταβλητή (y) είναι η εξαρτημένη ενώ οι κύριες μεταβλητές είναι η πιθανότητα φυλάκισης (Prob) και ο μέσος χρόνος φυλάκισης σε πολιτειακές φυλακές (Time).

# Μεταβλητές



| | |
|---|---|
| M | Ποσοστό ανδρών ηλικίας 14-24 |
| So | Δείκτρια μεταβλητή Νότιας Πολιτείας |
| Ed | Μέσος χρόνος εκπαίδευσης |
| Po1 | Αστυνομικές δαπάνες (1960) |
| Po2 | Αστυνομικές δαπάνες (1959) |
| LF | Αναλογία συμμετοχής εργατικού δυναμικού |
| M.F | Αριθμός ανδρών ανά 1000 γυναίκες |
| Pop | Πληθυσμός Πολιτείας |
| NW | Αριθμός μη λευκών ανα 1000 πολίτες |

| | |
|---|---|
| U1 | Λόγος ανεργίας για κατοίκους αστικών κέντρων ηλικίας 14-24 |
| U2 | Λόγος ανεργίας για κατοίκους αστικών κέντρων ηλικίας 35-39 |
| GDP | Κατά κεφαλήν ΑΕΠ |
| Ineq | Ανισότητα Εισοδήματος |
| Prob | Πιθανότητα φυλάκισης |
| Time | Μέσος χρόνος φυλάκισης σε πολιτειακές φυλακές |
| y | Αναλογία κατά κεφαλή εγκλημάτων για μια συγκεκριμένη κατηγορία |

# Αποτελέσματα...

| | HRM | | Zellner | | |
|---|---|---|---|---|---|
| | $MC^3$ | Occam's Window | $MC^3$ | RJMCMC | Stepwise |
| M | 0.65 | 0.92 | 0.65 | 0.64 | * |
| So | 0.17 | 0.04 | 0.13 | 0.12 | |
| Ed | 0.85 | 0.99 | 0.80 | 0.79 | * |
| Po1 | 0.64 | 0.53 | 0.56 | 0.55 | |
| Po2 | 0.59 | 0.47 | 0.47 | 0.48 | |
| LF | 0.10 | 0.03 | 0.11 | 0.11 | |
| M.F | 0.10 | 0.07 | 0.11 | 0.11 | |
| Pop | 0.22 | 0.36 | 0.19 | 0.19 | * |

| | $MC^3$ | Occam's Window | $MC^3$ | RJMCMC | Stepwise |
|---|---|---|---|---|---|
| NW | 0.36 | 0.59 | 0.25 | 0.25 | * |
| U1 | 0.10 | 0.12 | 0.10 | 0.11 | |
| U2 | 0.25 | 0.64 | 0.24 | 0.24 | * |
| GDP | 0.22 | 0.18 | 0.12 | 0.12 | * |
| Ineq | 0.97 | 1.00 | 0.96 | 0.96 | * |
| Prob | 0.61 | 0.91 | 0.56 | 0.56 | * |

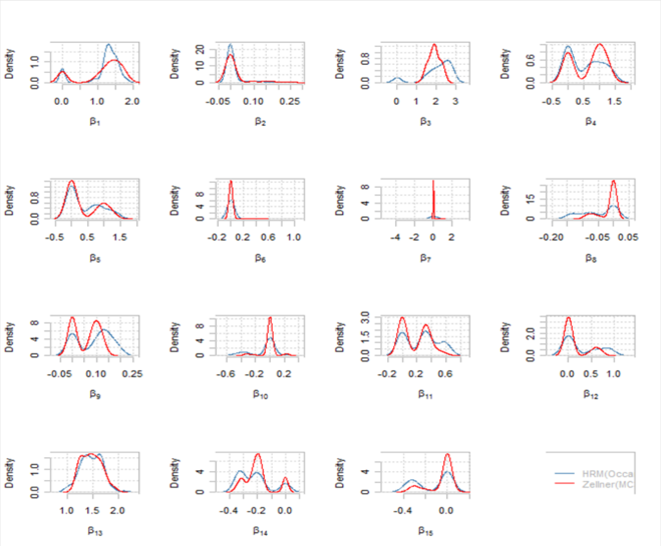
| | HRM | Zellner | Stepwise |
|---|---|---|---|
| RMSE | 0.187 | 0.231 | 0.253 |

| | HRM | | Zellner | |
|---|---|---|---|---|
| | $MC^3$ | Occam's Window | $MC^3$ | RJMCMC |
| $E(q\|y)$ | 6.97 | 6.29 | 6.80 | 6.78 |

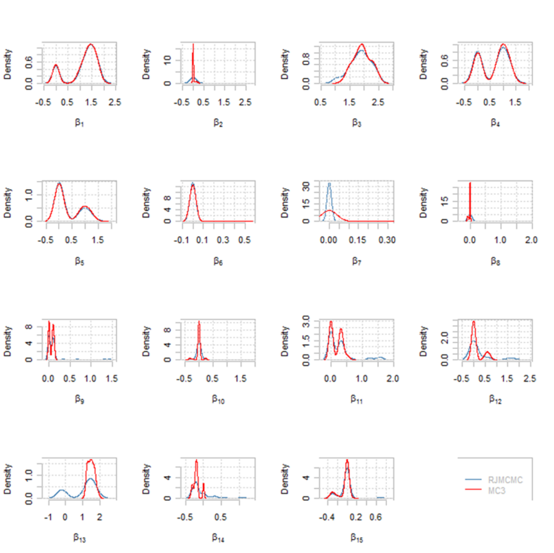
| | HRM | | Zellner | |
|---|---|---|---|---|
| | $MC^3$ | Occam's Window | $MC^3$ | RJMCMC |
| Επαναλήψεις | 800.000 | - | 800.000 | 800.000 |
| % Visits | 4.55 | - | 16.36 | 12.85 |
| Επιλεγέντα μοντέλα | - | 102 | - | - |
| $\rho$ | $> 0.95$ | - | $> 0.95$ | $> 0.95$ |

| | | M | So | Ed | Po1 | Po2 | LF | M.F |
|---|---|---|---|---|---|---|---|---|
| HRM (Occam's) | $E(\beta_i\|y)$ | 1.30 | 0.06 | 2.14 | 0.54 | 0.48 | 0.20 | -0.20 |
| | $\sqrt{Var(\beta_i\|y)}$ | 0.68 | 0.04 | 0.72 | 0.56 | 0.55 | 0.16 | 0.89 |
| Zellner ($MC^3$) | $E(\beta_i\|y)$ | 1.09 | 0.03 | 1.84 | 0.68 | 0.35 | 0.03 | 0.06 |
| | $\sqrt{Var(\beta_i\|y)}$ | 0.73 | 0.08 | 0.62 | 0.54 | 0.52 | 0.21 | 0.58 |
| Πλήρες* | $E(\beta_i\|y)$ | 2.52** | 0.14 | 2.14** | 0.55 | 0.22 | 1.27** | -4.03** |
| | | NW | U1 | U2 | GDP | Ineq | Prob | Time |
| HRM (Occam's ) | $E(\beta_i\|y)$ | 0.07 | -0.02 | 0.22 | 0.12 | 1.47 | -0.21 | -0.09 |
| | $\sqrt{Var(\beta_i\|y)}$ | 0.07 | 0.16 | 0.23 | 0.32 | 0.41 | 0.12 | 0.18 |
| Zellner (MC^{3}) | $E(\beta_i\|y)$ | 0.05 | -0.01 | 0.16 | 0.12 | 1.45 | -0.19 | -0.05 |
| | $\sqrt{Var(\beta_i\|y)}$ | 0.06 | 0.12 | 0.20 | 0.31 | 0.35 | 0.12 | 0.13 |
| Πλήρες* | $E(\beta_i\|y)$ | 0.05 | -0.09 | 0.46** | 0.71** | 1.68** | 0.45** | -0.73** |

# Εκ των υστέρων κατανομές των συντελεστών

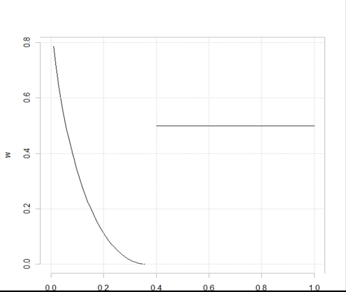


# $MC^3$ vs $rJ - MCMC$



Η προσέγγιση των Sellke *et al.*(2001) μπορεί να χρησιμοποιηθεί ως εκ των προτέρων πιθανότητα για την συμπερίληψη της $X_i$ σε ένα μοντέλο

$$w_i = \frac{1}{1 - e\Pr(|T| \geq t)\log\Pr(|T| \geq t)}, \ \Pr(|T| \geq t) \leq e^{-1}.$$



Θεωρώντας ότι η εκ των προτέρων κατανομή της $w$ είναι $Be(\tau,\lambda)$, $\tau,\lambda>0$ τότε η εκ των προτέρων κατανομή του μοντέλου $\gamma$ είναι

$$\pi(\gamma)=\binom{p}{q_\gamma}\frac{B(q_\gamma+\tau,p-q_\gamma+\lambda)}{B(\tau,\lambda)},\ q_\gamma=0,1,\ldots,p,$$

δηλαδή η $Bb(p,\tau,\lambda)$.

### $\tau,\lambda$ calibration

Για $\tau=1$ αρκεί να προσδιορίσουμε την παράμετρο $\lambda$. Βασιζόμενοι στο γεγονός ότι $E(\gamma)=\frac{\tau p}{(\tau+\lambda)}$ μπορούμε να γράφουμε $\lambda=\frac{p-\overline{q}}{\overline{q}}$ όπου $\overline{q}$ ο εκ των προτέρων αναμενόμενος αριθμός μεταβλητών που περιλαμβάνεται στο μοντέλο.

1. Στατιστική πρακτική
   - Επιλογή βάση Κριτηρίων Πληροφορίας
   - Προβλήματα διαδικασιών επιλογής
2. BMA...
   - Η βασική ιδέα
   - Βασικές Υλοποιήσεις
   - Αλγόριθμοι υλοποίησης
3. Εφαρμογή
4. Περαιτέρω θέματα στη BMA
   - Εναλλακτικός προσδιορισμός των $w_i$
   - Το $w$ ως τυχαία μεταβλητή
   - Εμπειρική εκτίμηση του $\phi$
   - Η κατανομή των Zellner-Siow

## Τοπικοί Εμπειρικοί εκτιμητές Bayes (τόσα $\phi$ όσα και μοντέλα...)

Ο εκτιμητής είναι ο

$$\widehat{\phi}_\gamma^{E(L)} = (F_\gamma - 1) \vee 0,$$

όπου $F_\gamma = \dfrac{R_\gamma^2/q_\gamma}{(1-R_\gamma^2)/(n-q_\gamma-1)}$ η συνήθης στατιστική συνάρτηση για τον έλεγχο της υπόθεσης $\beta_\gamma = 0$.

## Καθολικοί Εμπειρικοί εκτιμητές Bayes (μοναδικό $\phi$ )

Σε αυτή την περίπτωση ο εκτιμητής είναι

$$\widehat{\phi}_{\gamma}^{E(G)} = \arg\max_{\phi > 0} \sum_{\gamma} \left\{ \pi(\gamma) \frac{(1+\phi)^{(n-q_\gamma-1)}}{\left[\phi\left(R_\gamma^2 - 1\right) + 1\right]^{\left(\frac{n-1}{2}\right)}} \right\}.$$

Ο $\widehat{\phi}_{\gamma}^{E(G)}$ δεν δίνεται σε κλειστή μορφή...

### Λύση

Εφαρμογή του αλγόριθμου EM ακολουθώντας τους Liang *et al.* (2008).

Στην $i+1$ επανάληψη του αλγορίθμου

*E-Βήμα.* Υπολόγισε τις επόμενες αναμενόμενες τιμές για κάθε $\gamma \in \Gamma$

$$E(\sigma^2|y,\gamma,\widehat{\phi}_\gamma^{(i)}) = \frac{n-1}{SSY - \frac{\widehat{\phi}_\gamma^{(i)}}{1+\widehat{\phi}_\gamma^{(i)}} SSR},$$

$$E(\gamma|y,\widehat{\phi}_\gamma^{(i)}) = \frac{p(y|\gamma,\widehat{\phi}_\gamma^{(i)})}{\sum_\gamma p(y|\gamma,\widehat{\phi}_\gamma^{(i)})}.$$

*M-Βήμα.* Ο (τρέχων) εκτιμητής είναι

$$\widehat{\phi}_\gamma^{(i+1)} = \left[\frac{\sum_\gamma E(\gamma|y,\widehat{\phi}_\gamma^{(i)}) E(\sigma^2|y,\gamma,\widehat{\phi}_\gamma^{(i)}) SSR}{\sum_\gamma E(\gamma|y,\widehat{\phi}_\gamma^{(i)}) q_\gamma} - 1\right] \vee 0.$$

*Βήμα-ελέγχου.* Εάν $|\widehat{\phi}_\gamma^{(i+1)} - \widehat{\phi}_\gamma^{(i)}| < \varepsilon$ τότε $\widehat{\phi}_\gamma^{E(G)} = \widehat{\phi}_\gamma^{(i+1)}$.

## Information paradox

$$PO(\gamma;0) = c,\ c < \infty.$$

## Λύση του παραδόξου...

Οι Liang *et al.* (2008) χρησιμοποίησαν τις

$$\pi(\beta_\gamma|\sigma^2,\gamma) \propto \frac{\Gamma(q_\gamma/2)}{\pi^{q_\gamma/2}} \left|\frac{X_\gamma^T X_\gamma}{n\sigma^2}\right|^{1/2} \left(1 + \beta_\gamma^T \frac{X_\gamma^T X_\gamma}{n\sigma^2} \beta_\gamma\right)^{-q_\gamma/2},$$

$$\pi(\alpha,\sigma^2) \propto \sigma^{-2},\ \alpha \in \mathbb{R}\ \sigma > 0.$$

## Ιδιότητες της Z-S κατανομής

Για την εκ των προτέρων κατανομή των Zellner – Siow ισχύουν τα εξής :

- Είναι απαλλαγμένη από το «παράδοξο της πληροφορίας».
- Είναι συνεπής στην επιλογή του πραγματικού μοντέλου,
- Ο BMA-εκτιμητής της πρόβλεψης υπό την εκ των προτέρων κατανομή των Zellner – Siow είναι συνεπής ως προς την πρόβλεψη.

| | HRM | Zellner | Stepwise | Z-S |
|---|---|---|---|---|
| RMSE | 0.1875916 | 0. 2312784 | 0.2537807 | 0.2665655 |

# Συμπεράσματα...

- Η BMA είναι πράγματι αποτελεσματικότερη της συνήθους διαδικασίας.
- Τα παραδοσιακά μοντέλα παραμένουν ισχυρά έναντι των νεότερων προσθηκών (4.*2*-*4*).

- Με τι δεν ασχοληθήκαμε...
  - ΓΓΜ
  - Χρονοσειρές
  - Δέντρα
  - *και πολλά άλλα*...

# Βιβλιογραφία I


Claeskens & Hjort
*Model Selection and Model Averaging.*
Cambridge University Press, 2008.

Liang, F., Paolo, R., Molina, G., Clyde, M.A. & Berger J.O.
Mixtures of g-Priors for Bayesian Variable Selection.
*Journal of the American Statistical Association.* 103, pp. 410-423. (2008)

Raftery, A.E., Madigan, D. & Hoeting, J.A.
Bayesian model averaging for regression models.
*Journal of the American Statistical Association.* 92, pp. 179-191. (1997)